HHS Public Access

Author manuscript

J Registry Manag. Author manuscript; available in PMC 2015 March 16.

Published in final edited form as:

J Registry Manag. 2014 ; 41(2): 72–76.

Relative Survival Analysis Using the Centers for Disease Control and Prevention's National Program of Cancer Registries Surveillance System Data, 2000-2007

Reda J Wilson[1], **A Blythe Ryerson**[1], **Kevin Zhang**[2], and **Xing Dong**[2]

[1] Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA 30341, U.S.A.

[2]**ICF International, Fairfax, VA 22031, U.S.A.**

Abstract

BACKGROUND—Cancer survival rates are important to evaluate cancer care, identify disease patterns, and to estimate the probability of death due to cancer. To date, survival rates have been calculated using other data sets with limited population coverage that may not be able to fully identify differences by treatment, geographic regions, and racial or ethnic groups. Data from the Centers for Disease Control and Prevention's (CDC) National Program of Cancer Registries (NPCR) have not previously been used to calculate relative survival rates within the United States.

METHODS—Data from CDC's November 2011 submission for 21 state population-based central cancer registries, representing 50% of the U.S. population, were included in this analysis. This paper presents relative survival rates for diagnosis years 2000-2007 with follow-up through 2008.

RESULTS—The relative survival rate for all cancers and races combined was 65.0%; 65.3% for males, 64.8% for females. Blacks had a lower relative survival rate than whites, except for lung and bronchus. For all cancers, the <45 age groups had the highest relative survival rates, except for black males.

DISCUSSION—For all cancer primary sites combined for 2000-2007, the CDC NPCR five-year relative survival rate is comparable to that reported by the National Cancer Institute and the Canadian Cancer Registry. This analysis presents, for the first time, relative survival rates for half of the total U.S. population and demonstrates that reliable survival rates can be calculated using CDC's NPCR data now and in the future.

Keywords

Cancer Surveillance; Relative Survival Rates; NPCR; NPCR Cancer Surveillance System; Population-based Central Cancer Registries

Corresponding Author: Reda Wilson, MPH, CTR, RHIT Cancer Surveillance Branch Division of Cancer Prevention and Control National Center for Chronic Disease Prevention and Health Promotion Centers for Disease Control and Prevention 4770 Buford Highway, MS F-76 Atlanta, GA 30341 Office: 770-488-3245 Cell: 770-596-5896 FAX: 770-488-4759 dfo8@cdc.gov.

The authors of this report declare no conflict of interest in its development or presentation.

Introduction

Other publications [1-7] have shown the importance of cancer survival rates to evaluate cancer care, identify disease patterns, and to estimate the probability of death due to cancer. This information is of interest to clinicians, patients, public health practitioners, and researchers as well as the public. To date, survival rates have been calculated using other data sets with limited population coverage that may not be able to fully identify differences by treatment, geographic regions, and racial and ethnic groups [1].

Estimates of current U.S. survival rates are based on data from the National Cancer Institute (NCI) Surveillance, Epidemiology and End Results (SEER) program or from special studies [1, 8-10]. The NCI data set covers approximately 14% or 28% of the U.S. population, depending on the table referenced [1]. Until the preparation of this report, data from the Centers for Disease Control and Prevention's (CDC) National Program of Cancer Registries (NPCR) have not been used to calculate relative survival rates. This report demonstrates that researchers can also use NPCR data to calculate reliable relative survival rates that helps to provide a more comprehensive picture of cancer survival in the United States.

CDC's NPCR was established in 1992 to support the implementation and enhancement of population-based central cancer registries in U.S. states and territories [11]. CDC currently supports 45 states, the District of Columbia, Puerto Rico, and the Pacific Island Jurisdictions covering 96% of the U.S. population. CDC and NCI jointly support five of these, while the remaining five states are funded by NCI alone.

Beginning in 2001, NPCR-funded grantees began reporting incidence data to CDC [12]. Currently, CDC and NCI cancer surveillance data together are the foundation of national data on cancer trends including risks in special populations, the design and evaluation of public health interventions, many clinical and etiologic epidemiologic studies, and quality of care evaluation studies [13].

In order to understand cancer prognosis and improve control of all types of cancer in the United States, the survival rates for specific cancer sites and subpopulations may be monitored over time. Survival data are critical for evaluating the progress and impact of early detection and screening programs, comprehensive cancer control plans, and interventions from other sources [1-4]. As one available resource, survival data assists in identifying high-risk population groups which may lead to more effective and comprehensive program planning.

Methods

The relative survival rate is the ratio of the observed survival in a patient cohort divided by the expected survival of a comparable cancer-free group from the general population in the absence of the specific cancer [1]. A data set of invasive cancer cases for diagnosis years 2000-2007 from people ages 20-99 years was created using the NPCR November 2011 data submission file. We excluded data from 23 registries that either did not meet NPCR data quality standards for inclusion in the *United States Cancer Statistics* [14] or did not conduct active case follow-up or linkage with the National Center for Health Statistics' National

Death Index (NDI) through the 2008 death file. In 2006, the NPCR entered into an intra-agency agreement with NDI allowing registries to link with that database on a regular basis. An additional 4 registries requested their data not be used in survival analyses at this time. The quality and completeness of individual data items used in this analysis were previously conducted and reported elsewhere [15]. Due to difficulties with data collection following Hurricanes Katrina and Rita in 2005, we excluded data for July 1 through December 31, 2005 for Alabama, Louisiana, Mississippi, and Texas, according to established practices [1]. Twenty-one NPCR registries were included in the analysis representing 50% of the U.S. population.

Where incomplete, the full date of diagnosis and date of last contact were imputed [16]. Additionally, cases not linking to the state death files and NDI were presumed to be alive and the date of last contact was set to December 31, 2008. Cases with multiple primary cancers were included in the dataset, though only the first primary was included in calculating survival rates for all cancer sites combined. Where a patient had multiple primaries of the same site only the first primary was included in the calculations for that primary site [17-18]. We excluded cancer cases with unknown age or sex (1,873; 0.03%) and those identified solely on the basis of a death certificate or autopsy (75,253; 1%).

Using state-specific life tables by black and white race and sex, provided by the National Center for Health Statistics, [19] we calculated 5-year survival rates for cases diagnosed in 2000-2007 with follow-up through 2008 following a process described by Paul Dickman [2] that includes the Ederer II relative survival methodology. Rates were calculated for all cancer primary sites combined and 23 specific primary sites, grouped by the SEER Site Recode [1], by survival interval, sex, race, age group, and SEER Summary stage (local, regional, distant) [20]. The 23 specific primary site groups were selected due their frequency, screening amenability, and modifiable risk factors. All primary sites combined and the top ten primary site groups are shown in Table 1. Life tables are not available for races other than white and black populations, therefore survival rates are calculated for these races only. The age groups are defined as <45, 45-54, 55-64, 65-74, and 75 and older. All analyses were performed using SAS software version 9.3 (SAS Institute, Inc., Cary, North Carolina). Five-year relative survival rates are presented in Tables 1-3 for all cancer primary cancer sites combined and select specific primary sites by site, sex, age, and stage groups.

Results

More than five million cancer cases, from diagnosis years 2000-2007, were included in this dataset, and were evenly distributed between males and females. Cases for whites represented 85% of the total cases with 10% for blacks. The population from the 21 states was more than 300 million with a similar distribution for the white and black populations, 81% and 12% respectively (data not shown).

We present 5-year relative survival rates for all sites combined and specific cancer primary sites by sex and race in Table 1. The 5-year relative survival rate for all cancer primary sites combined, all races combined, was 65.0%; 65.3% for males and 64.8% for females. For all cancers and each specific cancer site, blacks had a lower 5-year relative survival rate than

whites, except lung and bronchus where the rate for black males was higher than white males and males of all races; 17.0%, 14.8%, and 14.7% respectively. Of the specific cancer sites analyzed, thyroid cancer had the highest relative 5-year survival rate for males and females combined (96.4%) and for females alone (97.6%). Prostate cancer had the highest relative 5-year survival rate for males (98.5%). The lowest rates were seen for lung and bronchus, ranging from 12.5% for black women to 20.4% for white women.

Table 2 shows 5-year relative survival rates by age at diagnosis and for all primary sites combined and the top four primary sites. For all cancer primary sites combined, the <45 age groups had the highest 5-year relative survival rates, except black males where the 65-74 age group had the highest rate. The 45-54 age groups had the highest relative survival rates for colon and rectum, except black males where the 65-74 age group had the highest rate. For lung and bronchus, the <45 age groups also had the highest relative survival rates. Age group 65-74 has the highest relative survival rates for prostate and female breast for all races combined, whites, and blacks.

Five-year relative survival rates by cancer stage at diagnosis are presented in Table 3. The highest rates are seen in the localized stage group and lowest rates in the distant stage group. This pattern is seen for all cancers combined and each specific cancer primary site analyzed. Blacks have lower 5-year relative survival rates than whites for all cancers combined, colon and rectum, lung and bronchus (except distant stage), and female breast. The relative survival rates for prostate for blacks are equal to or slightly higher than those seen for whites.

Discussion

This report provides 5-year relative survival rates for the largest proportion of the total U.S. population ever presented. While existing survival estimates cover 14% to 28% of the U.S. population[(1)], the NPCR data represents nearly double this population coverage and, when combined with NCI's data in the future, the two programs together, have the capability of covering 100% of the nation.

For all cancers combined for 2000-2007, the NPCR 5-year relative survival rate is 65.0%, comparable to that available from the NCI for the same time period, 67.1%, and the Canadian Cancer Registry, 62.0% for 2004-2006 [(3, 21)]. As expected, lower age groups have the highest survival as does localized stage at diagnosis. These trends are also comparable to those reported by NCI.

While SEER data have been a trusted and valued source of cancer survival information for decades, only a few states within the United States have previously been able to monitor their local cancer survival rates,[(16)] an essential tool in evaluating local cancer control initiatives. The addition of survival estimates from NPCR programs for national estimates will also provide more generalizable data, particularly for certain currently under-represented minority populations [(22)]. The NPCR program's expansive geographic coverage will also allow survival rates to be calculated for many rural populations and by geographic region. Although, through exclusions, the current report is not derived from the entire NPCR

program, the data reported here provide for the most complete description of the cancer burden in the United States to date [3]. Furthermore, the NPCR data set supports the analysis of rare cancers, such as retinoblastoma or primary fallopian tube cancers [23]. As the number of NPCR registries that link to the National Death Index increases, the quality and completeness of data on relative survival in the U.S. population will continue to grow as will our understanding of the impact of geographic, racial/ethnic, and medical care variations.

CDC's NPCR registries currently meet established standards for high quality data [14] and most, but not all, conduct either active case follow-up activities or linkages with the NDI in order to obtain and update vital status information. This means that NPCR data are now in a position where valid and reliable relative survival rates can be produced and published. However, additional work continues so that all NPCR registries will conduct data linkages with NDI on a regular basis. In addition to the inclusion requirements related to data quality and follow-up, NPCR grantees have the option to decline inclusion in the survival analyses at this time. CDC continues to encourage NPCR grantees to meet the USCS data quality and completeness criteria, to conduct follow-up data linkages, update their database following those linkages prior to the NPCR data submission, and allow inclusion of their data in these analyses.

CONCLUSION

The NPCR data are a robust data set available for calculating reliable 5-year relative survival rates that are not likely to be affected by small case counts. CDC will be including relative survival rates for a subset of NPCR programs in subsequent editions of the United States Cancer Statistics Web-Based Report [14]. The combination of survival estimates from both NCI and CDC will allow for more complete monitoring of any upcoming *Healthy People* objectives to increase the percentage of persons with cancer living five years or longer after diagnosis [24].

Acknowledgments

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

1. Howlader, N.; Noone, AM.; Krapcho, M.; Garshell, J.; Neyman, N.; Altekruse, SF.; Kosary, CL.; Yu, M.; Ruhl, J.; Tatalovich, Z.; Cho, H.; Mariotto, A.; Lewis, DR.; Chen, HS.; Feuer, EJ.; Cronin, KA., editors. SEER Cancer Statistics Review, 1975-2010. National Cancer Institute; Bethesda, MD: http://seer.cancer.gov/csr/1975_2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013
2. Dickman PW, Adami HO. Interpreting trends in cancer patient survival. Journal of Internal Medicine. 2006; 260:103–117. [PubMed: 16882274]
3. Ellison LF, Wilkins K. An update on cancer survival. Statistics Canada. 2010; 21(3) Catalogue no. 82-003-XPE.
4. Compton, CC., et al., editors. AJCC Cancer Staging Atlas: A Comparison to the Seventh Editions of the AJCC Cancer Staging Manual and Handbook. American Joint Committee on Cancer; 2012. DOI 10.1007/978-1-4614-2080-4_2

5. Allemani C, Rachet B, Weir HK, et al. Colorectal cancer survival in the USA and Europe: a CONCORD high-resolution study. BMJ Open. 2013; 3:e003055.
6. Allemani C, Sant M, Weir HK, et al. Breast cancer survival in the US and Europe: a CONCORD high-resolution study. Int J Cancer. 2013; 132:1170–1181. [PubMed: 22815141]
7. Coleman MP, Quaresman M, Berrino F, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). Lancet Oncol. 2008; 9:730–56. [PubMed: 18639491]
8. Abdollah F, et al. Incidence, survival and mortality rates of stage-specific bladder cancer in United States: a trend analysis. CancerEpidemiol. Jun; 2013 37(3):219–25.
9. Price GL, Davis KL, Karve S, Pohl G, Walgren RA. Survival Patterns in United States (US) Medicare Enrollees with Non-CML Myeloproliferative Neoplasms (MPN). PLoS ONE. 2014; 9(3):e90299. doi:10.1371/journal.pone.0090299. [PubMed: 24618579]
10. Schymura MJ, Kahn AR, German RR, Hsieh MC, Cress RD, Finch JL, Fulton JP, Shen T, Stuckart E. Factors associated with initial treatment and survival for clinically localized prostate cancer: results from the CDC-NPCR Patterns of Care Study (PoC1). BMC Cancer. 2010; 10:152. [PubMed: 20403178]
11. Hutton MD, Simpson LD, Miller DS, Weir HK, McDavid K, Hall HI. Progress Toward Nationwide Cancer Surveillance: An Evaluation of the National Program of Cancer Registries, 1994-1999. J Registry Manage. 2001; 28(3):113–120.
12. Thoburn KK, German RR, Lewis M, Nichols P, Ahmed F, Jackson-Thompson J. Case completeness and data accuracy in the Centers for Disease Control and Prevention's National Program of Cancer Registries. Cancer. 2007; 109:1607–1616. doi: 10.1002/cncr.22566. [PubMed: 17343277]
13. Division of Cancer Prevention and Control. Centers for Disease Control and Prevention; 2004/2005. Cancer Registries: The Foundation for Cancer Prevention and Control Fact Sheet..
14. U.S. Cancer Statistics Working Group. United States Cancer Statistics. 1999–2010 Incidence and Mortality Web-based Report. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; Atlanta: 2013. Available at: www.cdc.gov/uscs
15. Wilson R, O'Neil ME, Ntekop E, Zhang K, Ren Y. Coding Completeness and Quality of Relative Survival-Related Variables in the National Program of Cancer Registries Cancer Surveillance System 1995–2008. 2014 In progress.
16. Johnson CJ, Weir HK, Yin D, Niu X. The Impact of Patient Follow-up on Population-based Survival Rates. J Registry Manage. 2010; 37(3):86–103.
17. Ellison LF. Measuring the Effect of Including Multiple Cancers in Survival Analyses Using Data from the Canadian Cancer Registry. Cancer Epidemiology. 2010; 34:550–555. [PubMed: 20638928]
18. Brenner H, Hakulinen T. Patients with previous cancer should not be excluded in international comparative cancer survival studies. Int J Cancer. 2007; 121:2274–2278. [PubMed: 17594692]
19. Wei, R.; Anderson, RN.; Curtin, LR.; Arias, E. U.S. decennial life tables for 1999–2001: State life tables. National vital statistics reports. Vol. 60. National Center for Health Statistics; Hyattsville, MD: 2012.
20. Young, JL., Jr; Roffers, SD.; Ries, LAG.; Fritz, AG.; Hurlbut, AA., editors. SEER Summary Staging Manual - 2000: Codes and Coding Instructions. National Cancer Institute; Bethesda, MD: 2001. NIH Pub. No. 01-4969
21. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2011 Sub (1973-2009 varying) - Linked To County Attributes - Total U.S., 1969-2010 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2012, based on the November 2011 submission.
22. Merrill RM, Dearden KA. How Representative are the Surveillance, Epidemiology, and End Results (SEER) Program Cancer Data of the United States? Cancer Causes Control. 2005; 15(10): 1027–1034. [PubMed: 15801487]
23. Stewart SL, Wike JM, Foster SL, Michaud F. The incidence of primary fallopian tube cancer in the United States. Gynecol Oncol. Dec; 2007 107(3):392–7. Epub 2007 Oct 24. [PubMed: 17961642]

24. US Department of Health and Human Services. Health people 2020. US Department of Health and Human Services; Washington, DC: 2011. Available at http://www.healthypeople.gov/2020/topicsobjectives2020/default.aspx

Table 1

CDC NPCR-CSS[a] 5-Year Relative Survival (Percent) by SEER Cancer Primary Site Recode[b], Race[c], and Sex, 2000-2007

	All Races			Whites			Blacks		
Cancer Site	**Both Sexes**	**Males**	**Females**	**Both Sexes**	**Males**	**Females**	**Both Sexes**	**Males**	**Females**
All Sites Combined	65.0	65.3	64.8	65.8	65.8	65.9	59.3	62.0	56.2
Breast (female)	-	-	88.3	-	-	89.5	-	-	77.5
Colon & rectum	63.9	63.8	63.9	64.3	64.2	64.3	57.0	56.1	57.7
Corpus & uterus	-	-	80.7	-	-	82.8	-	-	58.9
Kidney & renal pelvis	69.2	69.0	69.4	69.1	69.0	69.3	67.3	67.0	67.8
Lung & bronchus	17.3	14.7	20.3	17.4	14.8	20.4	14.3	17.0	12.5
Melanoma of skin[d]	89.6	87.4	92.4	89.2	87.0	92.1	~	~	~
Non-Hodgkin lymphoma	65.7	64.0	67.6	66.0	64.5	67.8	58.1	54.7	62.1
Prostate	-	98.5	-	-	98.5	-	-	95.7	-
Thyroid	96.4	92.6	97.6	96.5	92.4	97.8	94.0	88.2	95.2
Urinary bladder	75.9	77.3	71.6	76.2	77.3	72.7	63.0	68.1	53.6

[a] Centers for Disease Control and Prevention National Program of Cancer Registries – Cancer Surveillance System

[b] Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2010, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975 2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.

[c] Life tables are not available to calculate survival rates on populations other than Whites and Blacks

[d] Melanoma of skin survival rates are limited to Whites.

Table 2

CDC NPCR-CSS[a] 5-Year Relative Survival (Percent) by Age at Diagnosis, SEER Cancer Primary Site Recode[b], Race, and Sex, 2000-2007

	Age at Diagnosis	Whites		Blacks	
		Males	Females	Males	Females
All Sites Combined	**Ages <45**	76.5	84.2	59.9	71.8
	Ages 45-54	66.7	78.0	58.1	63.7
	Ages 55-64	68.4	69.8	65.0	57.1
	Ages 65-74	67.8	61.2	67.6	50.8
	Ages 75+	55.7	49.4	51.8	38.9
Colon and Rectum	**Ages <45**	67.1	71.0	58.5	60.8
	Ages 45-54	69.1	71.2	59.1	64.6
	Ages 55-64	67.7	69.8	58.3	61.3
	Ages 65-74	66.1	67.3	59.2	60.6
	Ages 75-99	57.7	58.3	45.6	47.8
Lung and Bronchus	**Ages <45**	23.8	29.2	19.1	22.7
	Ages 45-54	17.0	26.0	13.4	19.4
	Ages 55-64	16.5	24.1	13.5	19.0
	Ages 65-74	15.6	21.5	12.5	16.7
	Ages 75-99	11.0	14.4	8.6	11.9

		Whites	Blacks
Prostate	**Ages <45**	95.1	96.4
	Ages 45-54	97.7	96.7
	Ages 55-64	98.9	97.3
	Ages 65-74	100.0	98.5
	Ages 75-99	93.1	84.4
Female Breast	**Ages <45**	88.0	76.3
	Ages 45-54	90.4	78.1
	Ages 55-64	90.2	78.6
	Ages 65-74	91.0	80.3
	Ages 75-99	87.2	73.0

[a] Centers for Disease Control and Prevention National Program of Cancer Registries – Cancer Surveillance System

[b] Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2010, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975 2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.

Table 3

CDC NPCR-CSS[a] 5-Year Relative Survival (Percent) by SEER Cancer Primary Site Recode[b], SEER Summary Stage at Diagnosis, Race, and Sex, 2000-2007

	SEER Summary Stage at Diagnosis	Whites		Blacks	
		Males	Females	Males	Females
All Sites Combined	**Localized**	91.7	91.0	92.2	84.6
	Regional	57.0	66.5	51.5	58.5
	Distant	24.8	25.7	20.3	21.7
Colon and Rectum	**Localized**	88.1	88.7	82.5	84.9
	Regional	67.8	68.5	63.7	63.6
	Distant	11.2	13.0	9.4	10.5
Lung and Bronchus	**Localized**	45.1	55.5	40.1	49.9
	Regional	19.6	25.5	17.9	23.1
	Distant	3.2	4.5	3.7	4.6

		Whites	Blacks
Prostate	**Localized**	100.0	100.0
	Regional	98.6	98.6
	Distant	28.8	30.9
Female Breast	**Localized**	98.1	92.3
	Regional	84.1	72.2
	Distant	25.6	16.8

[a] Centers for Disease Control and Prevention National Program of Cancer Registries – Cancer Surveillance System

[b] Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2010, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975 2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.